

AD-A182 753

THE EFFECTS OF INCREASING INFORMATION PROCESSING
DEMANDS ON RATING OUTCOMES(U) TEXAS A AND M UNIV
COLLEGE STATION R J FOTI ET AL OCT 86 TR-ONR-6
N88014-85-K-0289

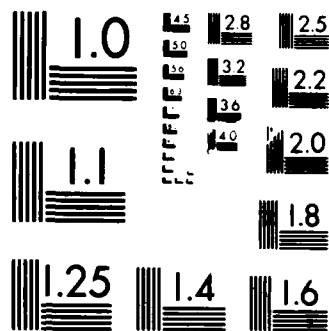
1/1

UNCLASSIFIED

F/G 5/8

NL

					END							
					8-87							
					DTIC							



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

AD-A182 753

12

DTIC FILE COPY

Human Resources Research

THE EFFECTS OF INCREASING INFORMATION
PROCESSING DEMANDS ON RATING OUTCOMES

Roseanne J. Foti
and
Neil M.A. Hauenstein

October, 1986
TR-ONR-6

Texas A&M University
and
Virginia Polytechnic Institute

DTIC
ELECTE
JUL 21 1987
S & E

This document has been approved
for public release and sale; its
distribution is unlimited.

87 7 20 025

THE EFFECTS OF INCREASING INFORMATION
PROCESSING DEMANDS ON RATING OUTCOMES

Roseanne J. Foti
and
Neil M.A. Hauenstein

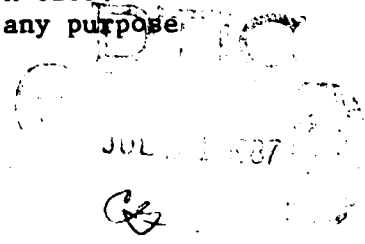
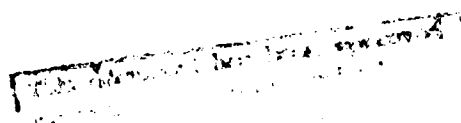
October, 1986
TR-ONR-6

Department of Psychology
Virginia Polytechnic Institute
and
State University

Prepared for:

Office of Naval Research
800 N. Quincy Street
Arlington, Virginia 22217

This report was prepared for the Manpower R&D Program of the
Office of Naval Research under contract N00014-85-k-0289.
Reproduction in whole or in part is permitted for any purpose
of the United States Government.



Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR-ONR-6	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER A182 753
4. TITLE (and Subtitle) The Effects Of Increasing Information Processing Demands On Rating Outcomes		5. TYPE OF REPORT & PERIOD COVERED Technical report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Roseanne J. Foti and Neil M.A. Hauenstein		8. CONTRACT OR GRANT NUMBER(s) N00014-85-k-0289
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Virginia Polytechnic Institute & State University Blacksburg, VA 24061		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-475-036
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of Navy Arlington, VA 22217		12. REPORT DATE October, 1986
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Supported by the Office of Naval Research Manpower R&D Program		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Performance appraisal, performance evaluation, rater error		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This research investigated the cognitive processes which mediate the performance rating process. Specifically, level of processing and rater prior performance information were manipulated in a 3 X 3 factorial design in order to assess the impact on psychometric rating outcomes and rating accuracy. Results indicated that as information processing demands increased, raters relied more on the past performance cues. Specifically, raters using more lenient ratings, and those using automatic processing and receiving a poor performance cue exhibited increased halo. In addition, raters were least accurate in recognizing behavior		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE
S N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

consistent with their performance cue. Implications for future research in performance appraisal are also discussed.

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Effects of Increasing Information Processing Demands on Rating Outcomes

Abstract

This research investigated the cognitive processes which mediate the performance rating process. Specifically, level of processing and ratee prior performance information were manipulated in a 3 X 3 factorial design in order to assess the impact on psychometric rating outcomes and rating accuracy. Results indicated that as information processing demands increased, raters relied more on the past performance cues. Specifically, raters using more automatic processing and receiving a good performance cue gave more lenient ratings, and those using automatic processing and receiving a poor performance cue exhibited increased halo. In addition, raters were least accurate in recognizing behaviors consistent with their performance cue. Implications for future research in performance appraisal are also discussed.

Report to be made by 10/15/75
Rating accuracy

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A-1	



Performance appraisal research is in a state of transition. For the last two decades, researchers have focused on basically three strategies for increasing rating validity: a) redesigning rating formats, b) training raters to minimize errors, and c) increasing observation skills. Empirical studies of the success of these approaches have shown some decrease in rating errors such as halo, but no corresponding increase in accuracy. In an effort to understand why these approaches have failed, and to gain greater insights into the determinants of rating accuracy, researchers have stressed the need for a new approach. This new approach focuses on analyzing the process underlying performance ratings (DeNisi, Cafferty, & Meglino, 1984, Feldman, 1981, Ilgen & Feldman, 1983).

Rating process research focuses on the rater's selection, storage, retrieval, and evaluation of information during the rating task. Many studies using this cognitive approach have appeared recently (for example, Banks, 1985; Murphy & Balzer, 1986; Murphy, Balzer, Lockhart, & Eisenman, 1985). In the typical rating process experiment, stimuli are presented in a manner that is relatively non-taxing of information processing capabilities. The sole task of the participants is to rate the performance of a target person (McIntyre, Smith, & Hassett, 1984; Murphy, Martin, & Garcia, 1982; Pulakos, 1984). In spite of the non-taxing nature of most experiments, recent research has shown that the processing of performance information can be biased by performance cues, by initial impressions, and by prior formal

judgments of the stimulus target. It is generally considered that these manipulations instill a schema that biases the rater's processing of subsequently seen information. (Balzer, 1986; DeNisi et al, 1984; Lord, 1985).

While these studies have demonstrated the potential of this cognitive approach for performance appraisal, Banks and Murphy (1985) have recently argued that this line of research is likely to widen the gap between research and practice. To prevent this from happening, it is essential for future research in performance appraisal to insure that the cognitive processes captured in the laboratory are similar to those in organizational settings. Thus, future laboratory studies should incorporate contextual variables, such as competing tasks, time pressures and delays between observing ratee behavior and appraisals, to enhance their external validity.

As noted by Feldman (1981), information about other organizational members generally occurs in complex and noisy informational environments. Supervisors are often simultaneously exposed to the behaviors of several individuals, complex task information and their own thoughts and memories. From these multiple sources, they must select relevant information and organize it into patterns that can be understood and remembered. Given our limited information capacity, selective attention in most organization environments is determined by automatic processes. Posner (1982), in reviewing work on attention and performance, notes that people can manage multiple tasks because they used automatic processes to simultaneously monitor several

informational sources with almost no interference. He points out that it is only when we "take notice of targets" (use controlled processes) that our capacity is severely limited and information sources interfere with one another.

During automatic processing, the information that is noticed is highly dependent on salient stimulus characteristics (e.g., loudness or uniqueness, Taylor & Fiske, 1978) as well as the cognitive schema guiding perception. In turn, the availability of schemata to guide information processing depends on primes such as past employee performance or the goals of the perceiver (Foti & Lord, 1987; Lord, 1985). There has been a debate in the current literature as to whether schematic processing causes a biased search for either confirming behaviors or inconsistent behaviors (Balzer, 1986; Murphy et al., 1985). However, under conditions which tax processing capabilities, it is more likely that raters will note schema inconsistent behaviors. For example, White and Carlston (1983) had subjects listen to two conversations simultaneously. Subjects were given prior personality information about one of the target stimuli. They found that subjects spent more time monitoring the conversation involving the target about which they had no prior information. In addition, subjects tended to switch their attention to the primed target's conversation when a schema inconsistent behavior occurred. These findings support Feldman's (1981) notion that supervisors will automatically process subordinates' performance until an atypical behavior causes the supervisor to move to a controlled level of processing.

The purpose of the present study was to assess the effect of increasing information processing demands on psychometric rating outcomes as well as rating accuracy. As processing demands increase, participants should rely more on automatic processing of performance. Since processing done automatically is highly influenced by the schema guiding perception, we hypothesized that performance cues presented prior to observation would have stronger impact as processing demands increased. More specifically, the present study utilized a 3 (level of processing) x 3 (good, poor or no performance information) design and predicted that as processing demands increased: (1) participants given a good performance cue would provide increasingly lenient ratings and those given a poor performance cue would provide increasingly strict ratings; (2) participants would rely more on prior performance cues, resulting in less differentiation of performance dimensions and thereby more halo; and (3) participants would be less accurate in recognizing behaviors consistent with their performance cue.

Method

Subjects

Participants in the study were 145 introductory psychology students, 73 males and 72 females, with a median age of 19 years. Participants were randomly assigned to experimental conditions with the stipulation that both males and females be equally distributed across conditions.

Stimulus Material

A 15-minute color videotape of an instructor lecturing on the topic area of consumer psychology served as the stimulus material. The tape was developed and used in a previous study by Hauenstein and Alexander (1986). Embedded in the tape were 16 behavioral incidents representative of four performance dimensions: four good behaviors representing the dimension organization, four good behaviors corresponding to depth of knowledge and two good and two poor behaviors for each dimension of delivery and relevance. This videotape was used because previous research indicated that it represented average performance, thereby avoiding a ceiling effect problem in testing the leniency/severity hypothesis.

Procedure

Participants reported to the lab in groups ranging from five to ten persons. All participants were told that they were about to watch a videotape of a brief lecture after which they would be asked to rate the performance of the instructor. Prior to viewing the videotape, participants were given written instructions containing the performance cue manipulation and the rating dimensions. After viewing the tape, participants completed a short filler task, the Picture-Number Test (Ekstrom, French, Harmon, & Derman, 1976) to eliminate the effects of short term memory, and then completed the rating form and a recognition memory questionnaire.

Experimental Manipulation

Level of processing. Three levels of this factor were manipulated, designed to create a continuum from controlled to automatic processing. In the edited condition, the instructor's behaviors were grouped (by editing the videotape) according to the four performance dimensions. This manipulation was designed to create extremely controlled processing because it eliminated for subjects the task of deciding which behaviors were most relevant for each performance dimension. In the tape condition, subjects were shown the videotape in normal order, similar to other laboratory studies of performance appraisal. Again, this manipulation was designed to create controlled processing although not as extreme as in the previous condition. In the task condition, participants were shown the videotape and also completed an additional task which involved thinking up at least 10 uses for two common objects. This manipulation was designed to create more automatic processing, since participants' attention was divided between the two tasks.

Performance cue. This factor was manipulated by written instructions. In each of the level of processing conditions described above, participants were either given a paragraph describing the past performance of the instructor as good, poor or they received no performance information.

Dependent Variables

Rating scales. The instructor's performance was evaluated using five 7-point graphic rating scales with anchors of poor and excellent. The five scales consisted of the four dimensions

embedded in the videotape and one dimension measuring an overall evaluation. Leniency was operationalized as simply the average dimension rating. Mean rating differences between the dimensions were expected because poor behaviors occurred on only two performance dimensions. Therefore, while halo was operationalized as participant's standard deviation across all four performance dimensions (Saal, Downey, & Lahey, 1980), subjects' ratings were converted to standard scores within each dimension prior to the computation of the halo scores (Pulakos, Schmitt, and Ostroff, 1986). The less the dispersion across the dimension, the smaller the standard deviation and the stronger the halo effect.

Recognition memory questionnaire. Recognition memory for whether specific instances of behavior were exhibited by the instructor was measured by a 32 item questionnaire. Eight items pertaining to each of the four dimensions represented on the videotape were included. Within each subset of eight items, four of the behaviors had appeared on the videotape, while the other four had not been exhibited by the instructor. Because more good than poor behaviors occurred in the stimulus tape, the questionnaire contained 24 good behaviors and 8 poor behaviors. Recognition accuracy for both good and poor behaviors was measured by the following formula:
$$\text{Number of true positives} + \text{number of true negative} \div \text{by the total number of behaviors.}$$
 True hits refer to the number of occurring behaviors correctly identified and true negatives refer to the number of non-occurring behaviors correctly identified.

Manipulation checks. A check for each experimental factor was included on a final questionnaire. Participants' perceptions of the amount of time their attention was focused on watching the instructor and the instructor's previous performance were each assessed by two 5-point Likert scale items.

Results

Manipulation Checks

A 3 (level of processing) X 3 (performance cue) analysis of variance (ANOVA) was used to assess the impact of the experimental manipulations on the participants' questionnaire responses. First, as expected, participants who were given an additional task to complete reported that they spent less time watching the instructor, ($\bar{M} = 2.35$), than did participants who only watched the tape, either in normal order ($\bar{M} = 3.11$) or edited ($\bar{M} = 3.59$), $F(2, 136) = 23.51$, $p < .001$. Second, participants in the good performance conditions rated their expectations for the instructor's performance significantly higher ($\bar{M} = 4.06$) than did participants in the poor ($\bar{M} = 1.15$) or no performance information ($\bar{M} = 2.92$) conditions, $F(2, 136) = 285.41$, $p < .001$.

Performance Ratings

Leniency. Since the experimental manipulations were successful, we can examine their effects on leniency. Hypothesis 1 predicted an interaction between level of processing and performance cue, such that participants using automatic processing (i.e., the task condition) and receiving the good performance cue would evaluate the instructor most leniently, and

those receiving the poor performance cue would evaluate the instructor more strictly. Results of a 3 X 3 multivariate analysis of variance (MANOVA) using all five performance dimensions as dependent variables provided strong support for this hypothesis, $F_{\text{approx}(40, 502)} = 2.45, p < .01$.

Insert Table 1 about here

As can be seen in the ANOVAS in Table 1, the level of processing X performance cue interaction was significant for three of the four individual dimensions as well as the overall rating. A priori comparisons were performed on the good versus poor cell means (see Table 2).

Insert Table 2 about here

The increasing magnitude of these deviations clearly showed that the performance cue manipulation had the least impact in the edited condition and the most impact in the task condition. Therefore, as predicted there was greater reliance on past performance information as information processing demands increased. Organization was the only dimension where this trend was not seen. The most likely explanation for this discrepancy was salience induced by the timing and behavioral content of this dimension. The critical behaviors for organization occurred either near the beginning or end of the videotape and typically

involved the lecturer moving from the lectern to the chalk board to refer to an outline of the lecture.

Halo. Hypothesis 2 predicted an ordinal interaction between level of processing and the performance cue due to the inclusion of a no performance cue condition in the design. Across all levels of processing demands, the subjects receiving no performance information were expected to sample both good and poor behaviors leading to greater differentiation across dimensions. For those subjects receiving prior performance information, the increase in processing demands would cause greater reliance on the performance cue leading to less differentiation across dimensions. A 3 (level of processing) X 3 (performance cue) ANOVA was performed on the halo scores. Only the level of processing main effect was significant, $F(2, 136) = 3.36$, $p < .05$. Halo tended to become stronger as processing demands increased across all levels of the performance cue condition (edited $M = .77$, tape $M = .68$, and task $M = .63$). These results suggest that there is a general tendency for halo strength to increase as processing demands during observation of performance increases, regardless of the schema used to aid the processing.

Recognition Memory

We expected that as information processing demands increased, participants would be least accurate in recognizing behaviors consistent with their performance cue. This would occur because raters using automatic processing would rely on a preexisting schema activated by the performance cue to process

the information and during retrieval would have trouble saying "no" to schema consistent behaviors that had not occurred.

Good behavior accuracy scores were entered in a 3 (level of processing) X 3 (type of performance cue) ANOVA. The expected interaction between level of processing and performance cue did not occur; instead two strong main effects were found (see Table 3).

Insert Table 3 about here

First, increasing processing demands caused a large decrease ($\eta^2 = .37$) in accuracy (edited $\bar{M} = .83$, tape $\bar{M} = .76$, task $\bar{M} = .69$). Second, the strong performance cue effect ($\eta^2 = .19$) was due to subjects receiving the good cue ($\bar{M} = .70$) being less accurate than subjects in poor cue ($\bar{M} = .78$) and no cue ($\bar{M} = .80$) conditions. As expected, examining the cell means in Table 3 clearly showed that the low true negative rates were causing the lower accuracy for subjects in the good cue condition. Subjects receiving the good cue were more likely to overestimate the frequency of good behaviors. The reason the interaction did not occur was because good cue subjects in the edited condition were just as likely to overestimate the frequency of good behavior as subjects in the more taxing tape and task conditions. Unlike the leniency results for the actual ratings, the performance cue biased recognition accuracy even in the least taxing condition.

Results for the recognition accuracy for poor behaviors were problematic. The same analysis as used for the good behaviors

resulted in no significant effects. As can be seen in Table 3, accuracy for poor performance was close to chance due to the low true negative rates. The commitment of many false positives could have been due to a strong negative impression. However, the performance ratings for the control group indicated an average impression. Closer examination of response patterns to the poor behavior foil items suggested that problem was due to the foil items. For example, one of the poor foil items for delivery was: Instructor used a lot of "uhms". Careful examination of the stimulus tape had shown the instructor as saying "uhm" only once. In the context of a fifteen minute lecture, apparently that was perceived as "a lot" of uhms. Three of the four poor foil items were subject to this type of problem.¹ Given the small total number of poor behaviors, this problem prevents an adequate test of our hypothesis.

Discussion

The present study was designed to examine the cognitive processes which mediate the performance rating process. Specifically, we manipulated the level of information processing demands on raters as well as varied ratee past performance information. Since processing done automatically is heavily influenced by the schema guiding perception, we expected that past performance information would be utilized more by raters as processing demands increased.

The results showed strong support for our first prediction concerning the leniency of evaluations. Subjects using automatic processing were influenced the most by prior performance

information. These results suggest that the initial categorization of the ratee by the supervisor is critical. Given that such consistent effects were found when using only one competing task, the effects of the supervisor's initial categorization would probably be much stronger under the typical high processing demands of day to day organizational life.

The results of the halo analysis suggest that under more automatic levels of processing raters demonstrate a stronger halo effect, regardless of the schema used to aid in processing the performance information. Murphy and Balzer (1986) found similar results when processing demands were taxed by a time delay. They suggested that systematic distortions by the raters caused the increase in halo. The most likely explanation of our finding is that increasing processing demands caused an undersampling of specific behaviors leading to increased reliance on the general impression (c.f., Cooper, 1981, p. 220). If there is a linear relationship (up to the point that processing demands are overwhelming) between processing demands and halo strength it raises questions concerning the notion that halo and accuracy are positively related (Cooper, 1981, Murphy & Balzer, 1986). It may be that this relationship does not generalize from the laboratory to the organization. If increasing processing demands increases halo strength, then halo in organizational performance appraisals could greatly overestimate the relationships among dimensions, possibly to the point that accuracy is no longer related to halo.

The recognition results for good behaviors raised some interesting issues. Even under the most nontaxing conditions,

subjects were less accurate at recognizing behaviors consistent with the performance cue. In contrast, for more categorical level judgments (i.e., performance ratings) the performance cue had little impact in the nontaxing condition. Several implications should be noted from this finding. Performance appraisal researchers should attempt to understand the differences in behavioral versus categorical judgments (c.f., Lord, 1985; Phillips & Lord, 1986). Schematic biases appear to be strongest at the behavioral level. Therefore, rater training programs that focus on the improving the quality of rating through accurate recall of specific behaviors are not likely to succeed (e.g., Thorton & Zorich, 1979). A better training strategy may be to focus on the supervisor's initial categorization of the ratee. To this end, frame-of-reference training (Bernardin & Buckley, 1981) may be the best vehicle. By training supervisors to adopt appropriate evaluative schemata, their critical initial categorization of the ratee should be more accurate (McIntyre et. al., 1982).

The finding that raters receiving the good performance cue were accurate recognizing good behaviors may also suggest how supervisors change their perceptions of workers. While the initial categorization of a worker is likely to be resistant to change (Fiske & Taylor, 1984), the supervisor is most likely to attend and store behaviors inconsistent with the initial categorization (c. f., White & Carlston, 1983). Our results are consistent with Feldman's notion that a worker exhibiting an unexpected behavior will cause elaborated processing on the part

of the supervisor. In terms of cognitive mechanisms governing this type of processing, the salience (Taylor & Fiske, 1978) of the unexpected behavior would attract the supervisors' attention and a schema-plus-tag model of memory (Graessar, Gordon, & Sawyer, 1979) could explain why it is more available in memory.

This type of cognitive process also has critical implications for the performance ratings in organizations. The major issue is how much contradictory information is needed before a supervisor can no longer discount the information and must recategorize (i.e., change their impression) the ratee. From a rater training perspective this issue is probably as important as the process involved in the initial categorization.

To summarize, the results of the present study advance our understanding of information processing in more realistic settings. Our results suggest that: 1) the schema utilized by the supervisor to form an initial impression of the ratee is critical; (2) after the initial categorization of the ratee, the supervisor is more likely to use controlled processing when attending to behaviors inconsistent with the initial impression; (3) the relationship between halo and accuracy may not generalize from the laboratory to the organization setting. It is important to remember that in the present study, raters' attention was divided between only two tasks. Future research will be able to determine if the patterns found in this study become more pronounced as information loads become heavier.

In conclusion, the results of the present research make it clear that future laboratory studies need to be more realistic.

If not, researchers cannot be certain that the cognitive processes we are tapping in the laboratory are similar to those utilized in organizational settings. Only by incorporating more contextual variables into future performance appraisal can we prevent the gap between research and practice from widening.

References

- Balzer, W. K. (1986). Biases in the recording of performance-related information: The effects of initial impression and centrality of the appraisal task. Organizational Behavior and Human Decision Processes, 37, 329-347.
- Banks, C. G. (1985, August). Cue selection and evaluation in performance appraisal. In R. L. Cardy (Chair), Information processing research in performance appraisal: A consideration of the approach, findings, and implications. Symposium conducted at the Academy of Management Association, San Diego.
- Banks, C. G., & Murphy, K. R. (1985). Toward narrowing the research-practice gap in performance appraisal. Personnel Psychology, 38, 335-345.
- Bernardin, H. J., & Buckley, M. R. (1981). Strategies in rater training, Academy of Management Review, 6, 205-212.
- Cooper, W. H. (1981). Ubiquitous halo. Psychological Bulletin, 90, 218-244.
- Denisi A., Cafferty, T. P., & Meglino, B. (1984). A cognitive view of the performance appraisal process: A model and research propositions. Organizational Behavior and Human Performance, 33, 360-396.
- Ekstrom, R. B., French, J. W., Harmon, H. H., & Derman, D. (1976). Manual for Kit of Factor-referenced Cognitive Tests. Princeton, NJ: Educational Testing Service.

- Feldman, J. M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66, 127-148.
- Fiske, S. T., & Taylor, S. E. (1984). Social cognition. Reading, MA, Addison-Wesley Publishing Company.
- Foti, R. J., & Lord, R. G. (1987). Prototypes and scripts: The effects of alternative methods of processing information on rating accuracy. Organizational Behavior and Human Decision Processes.
- Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer and tag hypothesis. Journal of Verbal Learning and Verbal Behavior, 18, 319-332.
- Hauenstein, N. M. A., & Alexander, R. A. (1986). Implicit personality theory as a cognitive difference factor in performance ratings. Unpublished Manuscript.
- Hauenstein, N. M. A., Whitcomb, A. J., & Foti, R. J. (1987). Effects of future interactions and time delays on evaluator leniency. Unpublished manuscript.
- Ilgen, D. R., & Feldman, J. M. (1983). Performance appraisal: A process focus. In B. Staw and L. L. Cumming (Eds.), Research in organizational behavior (Vol. 5). Greenwich, CN: JAI Press.

- Lord, R. G. (1985). Accuracy in behavioral measurement: An alternative definition based on rater's cognitive schema and signal detection theory. Journal of Applied Psychology, 70, 66-71.
- McIntyre, R., Smith, D., & Hassett, C. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of ratings. Journal of Applied Psychology, 69, 147-156.
- Murphy, K. R., Balzer, W. K., (1986). Systematic distortions in memory-based behavior ratings and performance evaluations: Consequences for rating accuracy. Journal of Applied Psychology. 71, 39-44.
- Murphy, K. R., Balzer, W., Lockhart, M., & Eisenman, E. (1985). Effects of previous performance on evaluations of present performance. Journal of Applied Psychology, 70, 72-84.
- Murphy, K. R., Martin, C., & Garcia, M. (1982). Do behavioral observation scales measure observation? Journal of Applied Psychology, 67, 562-567.
- Phillips, J., & Lord, R. G. (1982). Schematic information processing and perceptions of leadership in problem-solving groups. Journal of Applied Psychology, 67, 486-492.
- Posner, M. I. (1982). Cumulative development of attention theory. American Psychologist, 37, 168-179.
- Pulakos, E. D. (1984). A comparison of rater training programs: Error training and accuracy training. Journal of Applied Psychology, 69, 581-588.

- Pulakos, E. D., Schmitt, N., & Ostroff, C. (1986). A warning about the use of a standard deviation across dimensions within ratees to measure halo. Journal of Applied Psychology, 71, 29-32.
- Saal, F. E., Downey, R. G., & Lahey, M. A., Rating the Ratings: Assessing the psychometric quality of rating data. Psychological Bulletin, 1980, 88, 413-228.
- Taylor, S. E., & Fiske, S. T. (1978). Salience, attention, and attribution: Top of the head phenomenon. In L. Berkowitz (Ed.), Advances in experimental social psychology, (Vol. 11, pp. 249-288). Hillside, NJ: Erlbaum, 1981.
- White, D. J., & Carlston, D. E. (1983). Consequences of schema for attention, impressions, and recall of complex social interactions. Journal of Personality and Social Psychology, 45, 538-549.

Footnotes

- ¹ A recent study by Hauenstein, Whitcomb, & Foti (1987) was conducted using the same stimulus tape and a revised antiprototypical recognition measure. Pilot data for this study found much higher true negative rates for the antiprototypical foil items.

TABLE 1
Condition Means, Standard Deviations and Analyses of Variance for Leniency and Halo

Performance Dimensions ^a	Edited Tape		Tape Poor		Good		No		Task		Level of Processing (A)		Performance Que (B)		AXB F	
	Good	Good	Good	Good	Good	Good	No	No	Poor	No	F	F	F	F	F	F
Depth of Knowledge	4.73 (1.79)	3.57 (1.62)	3.80 (1.84)	4.93 (1.75)	3.18 (1.74)	4.65 (1.75)	2.47 (.91)	4.82 (2.01)	3.12 (1.82)	4.82 (2.01)	4.13*	5.09**	5.09**	5.09**	3.55**	3.55**
Delivery	3.27 (1.67)	2.93 (1.73)	2.71 (1.45)	3.93 (2.28)	2.24 (1.15)	2.47 (.91)	3.59 (2.06)	3.59 (2.06)	1.94 (.82)	3.59 (2.06)	6.11**	10.62***	10.62***	10.62***	6.77***	6.77***
Relevance	4.60 (1.55)	3.86 (1.66)	3.65 (1.41)	4.80 (1.52)	3.00 (1.32)	1.80 (.86)	4.29 (1.96)	4.29 (1.96)	3.12 (1.58)	4.29 (1.96)	9.65***	19.21***	19.21***	19.21***	2.45*	2.45*
Organization	5.56 (.86)	5.06 (1.68)	4.94 (1.39)	6.07 (1.10)	4.82 (1.33)	4.40 (1.12)	5.29 (1.69)	5.29 (1.69)	4.93 (1.54)	5.29 (1.69)	.91	8.33***	8.33***	8.33***	1.14	1.14
Overall	3.47 (1.68)	3.36 (1.73)	3.18 (1.42)	4.00 (1.36)	2.71 (1.16)	3.11 (.70)	3.88 (1.80)	3.88 (1.80)	2.59 (1.34)	3.88 (1.80)	8.86***	7.99***	7.99***	7.99***	3.71**	3.71**
Halo	.79 (.22)	.78 (.29)	.75 (.30)	.73 (.30)	.55 (.18)	.77 (.21)	.58 (.32)	.58 (.32)	.69 (.28)	.58 (.32)	3.36*	.28	.28	.28	1.65	1.65

Note. N = 145.

a. Higher values indicate more lenient ratings

b. Standard deviations appear in parentheses

*p < .05

**p < .01

***p < .001

Table 2

A priori Comparisons Between Performance Ratings in the Good
and Poor Performance Cue Conditions

	<u>Edited</u> <u>t</u>	<u>Tape</u> <u>t</u>	<u>Task</u> <u>t</u>
Depth of Knowledge	1.16*	1.75**	2.49***
Delivery	.34	1.69**	2.39***
Relevance	.74	1.80***	3.44***
Organization	.50	1.25**	1.20**
Overall	.19	1.29**	1.91***

Note. N = 145.

*p < .05

**p < .01

***p < .001

TABLE 3
Condition Means, Standard Deviations and Analyses of Variances for Recognition Accuracy

	Edited Tape		Tape		Task		Level of Processing (A)		Performance	
	Good	No	Good	No	Good	No	F	Que (B)	AXB F	
Good Behaviors										
True Hit Rate	.90 (.073) a.	.91 (.082)	.90 (.096)	.86 (.123)	.79 (.139)	.73 (.118)	.82 (.120)			
True Negative Rate	.69 (.076)	.82 (.126)	.53 (.108)	.71 (.082)	.45 (.082)	.72 (.104)	.67 (.120)			
Accuracy	.79 (.042)	.87 (.062)	.72 (.068)	.77 (.049)	.62 (.078)	.72 (.09)	.74 (.079)	52.85***	23.97***	1.06
Poor Behaviors										
True Hit Rate	.73 (.114)	.75 (.153)	.75 (.134)	.69 (.141)	.57 (.207)	.70 (.175)	.60 (.178)			
True Negative Rate	.35 (.296)	.25 (.234)	.38 (.208)	.31 (.226)	.58 (.284)	.48 (.229)	.48 (.312)			
Accuracy	.54 (.139)	.50 (.147)	.57 (.114)	.50 (.077)	.58 (.155)	.59 (.114)	.54 (.146)	2.78	2.10	.47

Note. N = 145.
aStandard deviation appear in parentheses
***p < .001

Manpower, Personnel, and Training R&D Program

Director of Research Programs
Office of Naval Research (Code 11)
Arlington, VA 22217-5000

Chairman, MPT R&D Planning Committee
Office of the Chief of Naval Research
Code 222
Arlington, VA 22217-5000

Life Sciences Technology Program
Manager (Code 125)
Office of the Chief of Naval Research
Arlington, VA 22217-5000

Defense Technical Information Center
DTIC/DDA-2
Cameron Station, Building 5
Alexandria, VA 22314

Science and Technology Division
Library of Congress
Washington, DC 20540

Office of the Assistant Secretary of
the Navy (Manpower & Reserve Affairs)
5D800, The Pentagon
Washington, DC 20350-1000

Team Head, Manpower, Personnel, and
Training Section
Office of the CNO (Op-914D)
4A578, The Pentagon
Washington, DC 20350-1000

Assistant for Research, Development
and Studies
Office of the DNCO(MPT) (Op-01B7)
Department of the Navy
Washington, DC 20370

Headquarters U.S. Marine Corps
Code MPI-20
Washington, DC 20380

Head, Leadership & Command
Effectiveness Branch (N-62F)
Naval Military Personnel Command
Department of the Navy
Washington, DC 20370-5620

Head, Human Factors Laboratory
Naval Training Systems Center (Code 71)
Orlando, FL 32813-7100

Technical Director
NPRDC (Code 01)
San Diego, CA 92152-6800

Director, Manpower and Personnel
Laboratory
NPRDC (Code 06)
San Diego, CA 92152-6800

Department of Administrative Sciences
Naval Postgraduate School (Code 54Ea)
Monterey, CA 93943-5100

Program Director
Manpower Research & Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Staff Specialist for Training
and Personnel Systems Technology
Office of the Under Secretary of
Defense for Research and Engineering
3D129, The Pentagon
Washington, DC 20301-3080

Technical Director
U.S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Benjamin Schneider
Department of Psychology
University of Maryland
College Park, MD 20742

Dr. Albert S. Glickman
Department of Psychology
Old Dominion University
Norfolk, VA 23508

Prof. Bernard M. Bass
School of Management
University Center at Binghamton
State U. of New York
Binghamton, NY 13901

Library
Naval War College
Newport, RI 02940

Page 2

Manpower, Personnel, and Training R&D Program

Lt. Col. Les Petty
NMCE
Headquarters, USMC
Washington, DC 20380

Col. Hester
NMPE
Headquarters, USMC
Washington, DC 20380

END

8-87

DTIC